

Limiting distribution of visits of several rotations to shrinking intervals

Ilya Vinogradov

May 11, 2010

Abstract

We show that given n normalized intervals on the unit circle, the numbers of visits of d random rotations to these intervals have a joint limiting distribution as lengths of trajectories tend to infinity. If d then tends to infinity, then the numbers of points in different intervals become asymptotically independent unless an arithmetic obstruction arises. This is a generalization of earlier results of J. Marklof.

The following question arises from two results of Marklof about gap distribution for rotations. Fix a point ξ in $[0, 1)$ and let $B_N = (0, N^{1/(d-1)}]^{\oplus(d-1)} \oplus \mathbf{R} \subset \mathbf{R}^d$. What is the limiting behavior of the number of points of the form

$$\left\{ \sum_{i=1}^{d-1} m_i \alpha_i \pmod{1} : m_i \in [1, N^{1/(d-1)}] \cap \mathbf{Z}, 1 \leq i \leq d-1 \right\},$$

for $\alpha_i \in [0, 1)$ that land in $(\xi - \frac{\sigma}{N}, \xi + \frac{\sigma}{N})$ as $N \rightarrow \infty$? In [1], J. Marklof showed that

$$\text{leb} \left\{ (\alpha, \xi) \in [0, 1)^{d-1} \times [0, 1) : \#\{\mathbf{m} \in B_N \cap \mathbf{Z}^d : \xi + \sum_{i=1}^{d-1} m_i \alpha_i + m_d \in (-\frac{\sigma}{N}, \frac{\sigma}{N})\} = A \right\} \rightarrow P^{(d)}(A)$$

as $N \rightarrow \infty$ and found its decay as $A \rightarrow \infty$. His main tool was the mixing property of a diagonal flow on $\text{SL}(d, \mathbf{R})/\text{SL}(d, \mathbf{Z})$ that had been proved by Moore [3]. In a later note Marklof remarked that for one variable (that is, $d = 2$), a stronger result is true due to a Theorem of Shah [6]. Namely, for fixed $\xi \in [0, 1) \setminus \mathbf{Q}$,

$$\text{leb} \left\{ \alpha \in [0, 1) : \#\{m \in \{1, \dots, N\} : \xi + m\alpha \pmod{1} \in (-\frac{\sigma}{N}, \frac{\sigma}{N})\} = A \right\} \rightarrow P^{(2)}(A).$$

This result uses Ratner's Theorem on measures invariant under unipotent flows [4]. We will generalize the theorems mentioned above to joint limiting probability distributions for several intervals and study their large d limits.

1 Notation and results

We will use the following notation.

- N, n , and d denote positive integers with $d \geq 2$;
- upper indices (usually j) run from 1 to n and lower indices (usually i) run from 1 to d unless stated otherwise;
- $\mathbf{m} = (m_1, \dots, m_d)$ is a vector of d integers;
- $\boldsymbol{\sigma} = (\sigma^1, \dots, \sigma^n)$ is a positive vector ($\sigma^j > 0$ for all j);
- $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_{d-1}) \in (\mathbf{R}/\mathbf{Z})^{d-1}$;
- $\boldsymbol{\xi} = (\xi^1, \dots, \xi^n) \in (\mathbf{R}/\mathbf{Z})^n$;
- $\boldsymbol{\tau} = (\tau^1, \dots, \tau^n)$ is a real vector;
- $\text{Pois } \sigma$ denotes the Poisson distribution with parameter σ .

If $n = 1$, we will write σ instead of σ^1 and $\boldsymbol{\sigma}$ and similarly for other variables. Let $B_N = (0, N^{1/(d-1)}] \oplus \mathbf{R} \subset \mathbf{R}^d$ as before. For a measurable set S define random variables $X_{\xi, S}^{N, d}: [0, 1) \rightarrow \mathbf{Z}$ by

$$X_{\xi, S}^{N, d} = \# \left\{ \mathbf{m} \in \mathbf{Z}^d \cap B_N : \sum_{i=1}^{d-1} m_i \alpha_i + m_d \in \xi + \frac{S}{N} \right\}.$$

We will usually suppress the upper indices on $X_{\xi, S}$. For a vector $\boldsymbol{\xi} \in \mathbf{T}^n$, the set $\Xi \subset \mathbf{T}^n$ is the closure of the orbit of rotation by $\boldsymbol{\xi}$ on the torus: $\Xi = \overline{\{k\boldsymbol{\xi} : k \in \mathbf{Z}\}}$; it is the smallest closed Lie subgroup of \mathbf{T}^n that contains $\boldsymbol{\xi}$.

Our results for limiting distributions of $X_{\xi, S}$ are as follows.

Theorem 1. *Fix any absolutely continuous probability measure on $[0, 1)$. With notation as above, the distribution of*

$$\mathbf{X}_{\boldsymbol{\xi}, \boldsymbol{\tau}, \boldsymbol{\sigma}} = (X_{\xi^1, (\tau^1, \tau^1 + \sigma^1)}, \dots, X_{\xi^n, (\tau^n, \tau^n + \sigma^n)})$$

has a weak limit as $N \rightarrow \infty$; we denote it by $\mathbf{P}_{n, \boldsymbol{\sigma}, \Xi, \boldsymbol{\tau}/\Xi}^{(d)}$.

In other words, the numbers of points in shrinking segments $(\xi^j + \frac{\tau^j}{N}, \xi^j + \frac{\sigma^j + \tau^j}{N})$, $1 \leq j \leq n$, with fixed “centers” ξ^j have a joint limiting distribution as N tends to infinity. The limiting distribution depends on d , n , $\boldsymbol{\sigma}$, Ξ , and $\boldsymbol{\tau}$ modulo Ξ . In particular, if $\Xi = \mathbf{T}^n$, then the distribution is independent of $\boldsymbol{\tau}$.

Remark 1. Jens Marklof proved special cases of this theorem. He proved the case $n = 1$, $d = 2$ in [2] and the case $n = 1$ and arbitrary d with average over ξ in [1].

Theorem 2. Let $\mathbf{P}_{n,\sigma,\Xi,\tau/\Xi}^{(d)}$ be the distribution from Theorem 1. Then, $\mathbf{P}_{n,\sigma,\Xi,\tau/\Xi}^{(d)}$ has a weak limit as $d \rightarrow \infty$. Furthermore,

$$\mathbf{P}_{n,\sigma,\Xi,\tau/\Xi}^{(d)} \implies (\text{Pois } \sigma_1, \dots, \text{Pois } \sigma_n)$$

as $d \rightarrow \infty$ iff

$$(\tau^j, \tau^j + \sigma^j) \cap (\tau^{j'}, \tau^{j'} + \sigma^{j'}) = \emptyset \text{ whenever } \xi^j = \xi^{j'}.$$

In effect, this Theorem says that as the number of rotations tends to infinity, the gap lengths exhibit random behavior. However, for every finite d and Ξ (with $n \geq 2$) we have that $\mathbf{P}_{n,\sigma,\Xi,\tau/\Xi}^{(d)}$ is dependent.

Acknowledgements. I would like to thank Ya. Sinai, E. Lindenstrauss, J. Marklof, F. Cellarosi, A. Salehi Golsefidy, and Z. Wang for helpful discussions.

2 Large N limit

Proof of Theorem 1. We reformulate the problem in the language of homogeneous spaces. Let $L = \text{SL}(d, \mathbf{R}) \ltimes (\mathbf{R}^d)^{\oplus n}$ and let $\Lambda = \text{SL}(d, \mathbf{Z}) \ltimes (\mathbf{Z}^d)^{\oplus n} \subset L$. Multiplication law on L is given by

$$(\mathbf{M}, \mathbf{v}^1, \dots, \mathbf{v}^n)(\mathbf{N}, \mathbf{w}^1, \dots, \mathbf{w}^n) = (\mathbf{MN}, \mathbf{v}^1 + \mathbf{M}\mathbf{w}^1, \dots, \mathbf{v}^n + \mathbf{M}\mathbf{w}^n).$$

It is well-known that $\Lambda \subset L$ is a non-cocompact lattice. The homogeneous space L/Λ is a bundle over $\text{SL}(d, \mathbf{R})/\text{SL}(d, \mathbf{Z})$ with fiber $(\mathbf{T}^d)^{\oplus n}$.

Given a set of vectors $\mathbf{v}^1, \dots, \mathbf{v}^n \in \mathbf{T}^d$, let

$$L_V = \{(1, \mathbf{v}^1, \dots, \mathbf{v}^n)^{-1}(\mathbf{M}, 0, \dots, 0)(1, \mathbf{v}^1, \dots, \mathbf{v}^n) \mid \mathbf{M} \in \text{SL}(d, \mathbf{R})\} \subset L;$$

it is of course isomorphic to $\text{SL}(d, \mathbf{R})$. Also define \hat{L}_V to be the smallest group containing L_V that is defined over \mathbf{Q} . Dimension of \hat{L}_V depends on \mathbf{v}^j . If all vectors \mathbf{v}^j have rational coordinates, then $\hat{L}_V = L_V$. Otherwise the fiber over the identity in \hat{L}_V is the smallest \mathbf{Q} -vector space containing the identity fiber for L_V . This construction can be carried to other points. Finally set $\hat{\Lambda}_V = \hat{L}_V \cap \Lambda$ which is a lattice in \hat{L}_V by construction.

For our purposes fix $\mathbf{v}^j = (0, \dots, 0, \xi^j)^T$. For this choice of \mathbf{v}^j we get the homogeneous space $\hat{L}_V/\hat{\Lambda}_V$. We have constructed \hat{L}_V so that $\pi(\hat{L}_V) = \overline{\pi(L_V)}$, where $\pi: L \rightarrow L/\Lambda$ is the canonical projection. The structure of this space depends on $\Xi = \overline{\mathbf{Z}\xi} \subset \mathbf{T}^n$. It is a subbundle of L/Λ : the base is still $\text{SL}(d, \mathbf{R})/\text{SL}(d, \mathbf{Z})$ but the fiber is Ξ^d after reordering coordinates.

We define $f_{\tau,\sigma}: \hat{L}_V/\hat{\Lambda}_V \rightarrow \mathbf{R}^n$ by

$$f_{\tau,\sigma}(\mathbf{M}, \mathbf{v}^1, \dots, \mathbf{v}^n) = (g_{\tau^1, \sigma^1}(\mathbf{M}, \mathbf{v}^1), \dots, g_{\tau^n, \sigma^n}(\mathbf{M}, \mathbf{v}^n)),$$

where

$$g_{\tau,\sigma}(\mathbf{M}, \mathbf{v}) = \sum_{\mathbf{m} \in \mathbf{Z}^d \setminus \{0\}} \chi_1(\tilde{m}_1) \dots \chi_1(\tilde{m}_{d-1}) \chi_{(\tau, \tau+\sigma)}(\tilde{m}_d),$$

$$\chi_\sigma(x) = \begin{cases} 1 & x \in (0, \sigma) \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\tilde{\mathbf{m}} = \mathbf{M}\mathbf{m} + \mathbf{v}.$$

It is easily seen that f is $\hat{\Lambda}_V$ invariant and hence well-defined on the quotient.

We need to show that $\text{leb}\{f_{\tau,\sigma} = (A^1, \dots, A^n)\} \rightarrow \mathbf{P}_{n,\sigma,\Xi,\tau/\Xi}^{(d)}(A^1, \dots, A^n)$ as $N \rightarrow \infty$. To this end we use a theorem of Shah (Theorem 1.4 in [6]). The form we need is the following:

Theorem 3 (Shah). *Let*

$$U_{\alpha} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ \alpha_1 & \dots & \alpha_{d-1} & 1 \end{pmatrix} \quad \text{and} \quad \Phi_t = \begin{pmatrix} e^{-t} & & & \\ & \ddots & & \\ & & e^{-t} & \\ & & & e^{(d-1)t} \end{pmatrix}.$$

Let L be a Lie group, $\Lambda \subset L$ a lattice, $\varphi: \text{SL}(d, \mathbf{R}) \rightarrow L$ an embedding. If the image of φ is dense when projected to L/Λ , then for any bounded continuous η

$$\lim_{t \rightarrow \infty} \int_{\mathbf{R}^{d-1}} \eta(\varphi(\Phi_t U_{\alpha})) d\nu(\alpha) = \int_{L/\Lambda} \eta(\mathbf{M}) d\mu(\mathbf{M}), \quad (1)$$

where ν is any absolutely continuous probability measure on U_{α} and μ is the Haar probability measure on L/Λ .

Remark 2. In effect, the Theorem says that the unstable manifold U_{α} is equidistributed in the larger homogeneous space L/Λ provided the density assumption is satisfied.

We set $N = e^{(d-1)t}$ and apply the Theorem 3 with $L = \hat{L}_V$, $\Lambda = \hat{\Lambda}_V$, and

$$\varphi: \mathbf{M} \mapsto (1, \mathbf{v}^1, \dots, \mathbf{v}^n)^{-1}(\mathbf{M}, 0, \dots, 0)(1, \mathbf{v}^1, \dots, \mathbf{v}^n).$$

This ensures density after projecting to $\hat{L}_V/\hat{\Lambda}_V$ by construction. We now use the following elementary Lemma to construct appropriate functions η .

Lemma 4. *Let $\mathbf{N}_0 = \mathbf{N} \cup \{0\}$. Then $p_2: \mathbf{N}_0 \times \mathbf{N}_0 \rightarrow \mathbf{N}_0$ given by*

$$(x, y) \mapsto \binom{x + y + 2}{2} - (y + 1)$$

is a bijection.

By induction, there exists a polynomial bijection between \mathbf{N}_0^n and \mathbf{N}_0 for each n ; call it p_n . Set

$$h_n(\mathbf{M}, \mathbf{x}^1, \dots, \mathbf{x}^n) = p_n(g_{\tau^1, \sigma^1}(\mathbf{M}, \mathbf{x}^1), \dots, g_{\tau^n, \sigma^n}(\mathbf{M}, \mathbf{x}^n))$$

and apply Shah's Theorem to functions

$$\eta_{\mathbf{A}}(\mathbf{N}, \mathbf{y}^1, \dots, \mathbf{y}^n) = \chi_{\{h_n(\mathbf{M}, \mathbf{x}^1, \dots, \mathbf{x}^n) = p_n(A^1, \dots, A^n)\}}((1, \mathbf{v}^1, \dots, \mathbf{v}^n)(\mathbf{N}, \mathbf{y}^1, \dots, \mathbf{y}^n))$$

for all nonnegative integers A^j . These functions are not continuous, but are indicators of nice sets. Using a standard approximation argument we can apply the Theorem to them as well. For \mathbf{M} of the form

$$\mathbf{M} = \begin{pmatrix} N^{-1/(d-1)} & & & \\ & \ddots & & \\ & & N^{-1/(d-1)} & \\ & & & N \end{pmatrix} \cdot \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ \alpha_1 & \dots & \alpha_{d-1} & 1 \end{pmatrix},$$

we recover the numbers of points in the n segments. In fact, for all \mathbf{M}

$$\eta_{\mathbf{A}}((1, \mathbf{v}^1, \dots, \mathbf{v}^n)^{-1}(\mathbf{M}, 0, \dots, 0)(1, \mathbf{v}^1, \dots, \mathbf{v}^n)) = \chi_{\{h_n(\mathbf{M}, \mathbf{x}^1, \dots, \mathbf{x}^n) = p_n(\mathbf{A})\}}(\mathbf{M}, \mathbf{M}\mathbf{v}^1, \dots, \mathbf{M}\mathbf{v}^n)$$

and

$$g_{\tau, \sigma}(\mathbf{M}, \mathbf{M}\mathbf{v}) = \sum_{\mathbf{m} \in \mathbf{Z}^d \setminus \{0\}} \chi_1 \dots \chi_1 \chi_{(\tau, \tau + \sigma)}(\mathbf{M}(\mathbf{m} + \mathbf{v})),$$

and for the particular choice of \mathbf{M} above,

$$\mathbf{M}(\mathbf{m} + \mathbf{v}) = \begin{pmatrix} m_1/N^{-1/(d-1)} \\ \vdots \\ m_{d-1}/N^{-1/(d-1)} \\ (m_1\alpha_1 + \dots + m_{d-1}\alpha_{d-1} + \xi)N \end{pmatrix}.$$

Thus, $\eta_{\mathbf{A}}(\varphi(\mathbf{M})) = 1$ if and only if there are exactly A^1, \dots, A^n visits to the segments around ξ^1, \dots, ξ^n for the given $\boldsymbol{\alpha}$ and N . Hence the form of the limiting distribution is

$$\mathbf{P}_{n, \sigma, \Xi, \tau/\Xi}^{(d)}(A^1, \dots, A^n) = \mu_{\hat{L}_V/\hat{\Lambda}_V}\{f_{\tau, \sigma}(\mathbf{M}, \mathbf{x}^1, \dots, \mathbf{x}^n) = (A^1, \dots, A^n)\}. \quad (2)$$

□

3 Large d Limit

In this section we consider the large d limit of the distributions from the previous sections and prove Theorem 2. Before proving the theorem, we will need basic information about the Poisson distribution.

Poisson distribution with parameter σ weighs each non-negative integer k with weight $e^{-\sigma}\sigma^k/k!$. We will denote Poisson distribution with parameter σ by $\text{Pois } \sigma$. Its moments have the form

$$\sum_{k=0}^{\infty} k^n e^{-\sigma} \frac{\sigma^k}{k!} = e^{-\sigma} \left(\sigma \frac{d}{d\sigma} \right)^n e^{\sigma} = \sum_{k=1}^n S(n, k) \sigma^k,$$

where $S(n, k)$ is the Stirling number of the second kind. As can be easily seen from the above equality, the Stirling number is the number of partitions of a set of n elements into k nonempty sets. The first few moments of the Poisson distribution are σ , $\sigma^2 + \sigma$, $\sigma^3 + 3\sigma^2 + \sigma$. These correspond to partitions $\{1\}$; $\{12\}$, $\{11\}$; $\{123\}$, $\{112\}$, $\{121\}$, $\{211\}$, $\{111\}$.

To further study the limiting distributions we have obtained, we will need the following generalization of a proposition of Marklof from [1] which goes back to a theorem of Rogers [5]. Let $\text{Gr}(n, l) = O(n)/(O(l) \times O(n-l))$ denote the Grassmannian of l -planes in \mathbf{R}^n ; we assume that the l -planes are embedded in \mathbf{R}^n with respect to the standard basis. Let $\text{Gr}(n, l)(\mathbf{Q}) = \{\pi \in \text{Gr}(n, l) \mid \pi \subset \mathbf{R}^n \text{ is defined over } \mathbf{Q}\} = \{\pi \in \text{Gr}(n, l) \mid \pi \cap \mathbf{Z}^n \text{ is a lattice in } \pi\}$. For $\pi \in \text{Gr}(n, l)(\mathbf{Q})$, we write $\text{covol } \pi_{\mathbf{Z}}$ for the covolume of the lattice $\pi_{\mathbf{Z}} = \pi \cap \mathbf{Z}^n$ in π . We also set $G = \text{SL}(d, \mathbf{R})$, $\Gamma = \text{SL}(d, \mathbf{Z})$, and fix μ to be the Haar probability measure on G/Γ .

Theorem 5. *Let $F: (\mathbf{R}^d)^{\oplus r} \rightarrow \mathbf{R}$ be a bounded piecewise continuous function with compact support. Let $f: G/\Gamma \rightarrow \mathbf{R}$ be defined by*

$$f(\mathbf{M}) = \sum_{\mathbf{m}^1, \dots, \mathbf{m}^r \in \mathbf{Z}^d} F(\mathbf{M}\mathbf{m}^1, \dots, \mathbf{M}\mathbf{m}^r)$$

with $r < d$ a positive integer. Then, the first moment of f is given by the following expression:

$$\int_{G/\Gamma} f(\mathbf{M}) d\mu(\mathbf{M}) = \sum_{l=0}^r \sum_{\pi \in \text{Gr}(r, l)(\mathbf{Q})} \int_{x \in \pi'} F(x) \frac{dx}{(\text{covol } \pi_{\mathbf{Z}})^d}, \quad (3)$$

where $\pi' \in \text{Gr}(rd, ld)(\mathbf{Q})$ is the image of π under the embedding

$$(x^1, \dots, x^r) \mapsto (x^1, \dots, x^1, \dots, x^r, \dots, x^r)$$

and the measure dx is the Lebesgue measure on π' that should be interpreted as the delta measure at the origin when $l = 0$.

Remark 3. If in the sum over \mathbf{m}^j we omit the terms where any of the \mathbf{m}^j are $\mathbf{0}$, then in the sum over $\pi \in \text{Gr}(r, l)(\mathbf{Q})$ we omit planes that are generated by subsets of the standard basis. This follows from the fact that such subsets of $(\mathbf{Z}^d)^r$ are $\text{SL}(d, \mathbf{Z})$ -invariant.

Lemma 6. *With notation as in the Theorem, we have*

$$\int_{G/\Gamma} \sum_{\substack{\mathbf{m}^1, \dots, \mathbf{m}^r \in \mathbf{Z}^d \\ \text{linearly indep.}}} F(\mathbf{M}\mathbf{m}^1, \dots, \mathbf{M}\mathbf{m}^r) d\mu(\mathbf{M}) = \int_{\mathbf{x}^j \in \mathbf{R}^d} F(\mathbf{x}^1, \dots, \mathbf{x}^r) d\mathbf{x}^1 \dots d\mathbf{x}^r.$$

Proof. First note that the integral is well-defined since linearly independent sets of vectors are preserved by Γ . Further renormalize μ so that $\mu(G/\Gamma) = \prod_{k=2}^d \zeta(k)$ for $d \geq 2$ and write $d\mu(\mathbf{M})/\mu(G/\Gamma)$ in the integral; this normalization will be useful later. Write

$$\mathbf{M} = \begin{pmatrix} x_{11} & \dots & x_{1d} \\ \vdots & \ddots & \vdots \\ x_{d1} & \dots & x_{dd} \end{pmatrix} \in G/\Gamma.$$

Then for $1 \leq r < d$ we have

$$\mathbf{M} = \left(\begin{array}{ccc|ccc} x_{11} & \cdots & x_{1r} & & & \\ \vdots & \ddots & \vdots & & & \\ x_{r1} & \cdots & x_{rr} & & & \\ \hline x_{r+1,1} & \cdots & x_{r+1,r} & \det^{-1}(x_{ij})_{i,j \leq r} & 0 & \cdots & 0 \\ x_{r+2,1} & \cdots & x_{r+2,r} & 0 & 1 & & \\ \vdots & \ddots & \vdots & \vdots & & \ddots & \\ x_{d1} & \cdots & x_{dr} & 0 & & & 1 \end{array} \right) \cdot \left(\begin{array}{ccc|ccc} & & & & & & \\ & & & & & & \\ & & & & & & \\ \hline & & & \text{Id}_{r \times r} & & & \\ & & & \vdots & \ddots & & \\ & & & z_{r1} & \cdots & z_{r,d-r} & \\ \hline & & & 0_{(d-r) \times r} & & & \\ & & & y_{11} & \cdots & y_{1,d-r} & \\ & & & \vdots & \ddots & & \\ & & & y_{d-r,1} & \cdots & y_{d-r,d-r} & \end{array} \right)$$

where $(y_{ij}) \in \text{SL}(d-r, \mathbf{R})$. In these coordinates

$$d\mu = \prod_{\substack{i \leq d \\ j \leq d-r}} dx_{ij} \prod_{\substack{i \leq r \\ j \leq d-r}} dz_{ij} \cdot \delta(1 - \det(y_{ij})) \prod_{i,j \leq d-r} dy_{ij}. \quad (4)$$

The last factor is the Haar measure on $\text{SL}(d-r, \mathbf{R})$ normalized to $\zeta(2) \cdots \zeta(d-r)$ (or simply 1 in case $d-r=1$).

For $j = 1, \dots, r$ let $t^j = \gcd \mathbf{m}^j$. Writing \mathbf{m}^j/t^j for a column vector, we can find $\mathbf{N} \in \text{SL}(d, \mathbf{Z})$ such that

$$\left(\frac{\mathbf{m}^1}{t^1} \quad \cdots \quad \frac{\mathbf{m}^r}{t^r} \right) = \mathbf{N} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ 0 & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{rr} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} = \mathbf{N}A.$$

A is a matrix with integer entries uniquely determined by the following conditions:

- a_{ij} , $1 \leq i \leq j$ are relatively prime for any fixed $j \in \{1, \dots, r\}$ (in particular, $a_{11} = 1$);
- $0 \leq a_{1j}, \dots, a_{j-1,j} < a_{jj}$.

The first condition is due to relative primality of \mathbf{m}^j/t^j , and the second comes from applying row operations. Given $a_{11} = 1, a_{22}, \dots, a_{rr}$, the number of possible matrices A of this form is

$$\prod_{j=1}^r \varphi_{j-1}(a_{jj}),$$

where φ_k is the number-theoretic function defined by

$$\varphi_k(p^\varepsilon) = p^{\varepsilon k} \left(1 - \frac{1}{p^k} \right)$$

for $k \geq 1$ and φ_0 is identically 1; φ_1 is Euler totient function. The function $\varphi_k(n)$ counts the number of k -tuples $(n_1, \dots, n_k) \in \{0, \dots, n-1\}^k$ such that $\gcd(n, n_1, \dots, n_k) = 1$.

Let us compute the stabilizer of a fixed matrix A :

$$\Gamma_A = \{\gamma \in \Gamma \mid \gamma A = A\} = \left(\begin{array}{c|c} \text{Id}_{r \times r} & \mathbf{Z}_{r \times (d-r)} \\ \hline 0_{(d-r) \times r} & \text{SL}(d-r, \mathbf{Z}) \end{array} \right).$$

Thus we get

$$\begin{aligned} \frac{1}{\mu(G/\Gamma)} \int_{G/\Gamma} \sum_{\mathbf{m}^j \text{ l. i.}} F(\mathbf{M}\mathbf{m}^1, \dots, \mathbf{M}\mathbf{m}^r) d\mu(\mathbf{M}) = \\ \frac{1}{\mu(G/\Gamma)} \sum_{t^1, \dots, t^r=1}^{\infty} \int_{G/\Gamma} \sum_{\mathbf{N} \in \Gamma/\Gamma_A} F(\mathbf{M}\mathbf{N} \begin{pmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{M}\mathbf{N} \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{M}\mathbf{N} \begin{pmatrix} a_{1r} \\ \vdots \\ a_{rr} \\ \vdots \end{pmatrix}) d\mu(\mathbf{M}) = \\ \frac{1}{\mu(G/\Gamma)} \sum_{t^j} \int_{G/\Gamma_A} F(\mathbf{M} \begin{pmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{M} \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{M} \begin{pmatrix} a_{1r} \\ \vdots \\ a_{rr} \\ \vdots \end{pmatrix}) d\mu(\mathbf{M}). \quad (5) \end{aligned}$$

Using (4) to change the measure we get

$$\frac{1}{\zeta(d) \dots \zeta(d-r+1)} \sum_{t^j} \sum_A \int_{(\mathbf{R}^d)^r} F(t^1 \mathbf{x}^1, t^2 a_{12} \mathbf{x}^1 + t^2 a_{22} \mathbf{x}^2, \dots, t^r a_{1r} \mathbf{x}^1 + \dots + t^r a_{rr} \mathbf{x}^r) d\mathbf{x}^1 \dots d\mathbf{x}^r.$$

Now we do a linear change of variables and get

$$\frac{1}{\zeta(d) \dots \zeta(d-r+1)} \sum_{t^j=1}^{\infty} \frac{1}{(t^1 \dots t^r)^d} \sum_{a_{22}, \dots, a_{rr}=1}^{\infty} \frac{\varphi_1(a_{22})}{a_{22}^d} \dots \frac{\varphi_{r-1}(a_{rr})}{a_{rr}^d} \cdot \int_{(\mathbf{R}^d)^r} F(\mathbf{x}^1, \dots, \mathbf{x}^r) d\mathbf{x}^1 \dots d\mathbf{x}^r. \quad (6)$$

It is easy to see that

$$\sum_{n \geq 1} \frac{\varphi_k(n)}{n^d} = \frac{\zeta(d-k)}{\zeta(d)},$$

whence the constant in front of the integral in (6) is 1, as desired. □

Proof of Theorem 5. Rewrite the integral we are evaluating as

$$\frac{1}{\mu(G/\Gamma)} \int_{G/\Gamma} \sum_{l=0}^r \sum_{\text{rk}(\mathbf{m}^1, \dots, \mathbf{m}^r)=l} F(\mathbf{M}\mathbf{m}^1, \dots, \mathbf{M}\mathbf{m}^r) d\mu(\mathbf{M}).$$

Here we normalize μ like in the Lemma and the inner sum runs over those r -tuples of vectors whose \mathbf{R} -span has dimension l . Since the set $\{\text{rk } \mathbf{m} = l\}$ is $\text{SL}(d, \mathbf{Z})$ -invariant for each l , we can pass the sum over l through the integral sign. To prove the Theorem, it suffices to show that corresponding terms in the expression above and in (3) match for each l . Now observe that

$$\{\text{rk } \mathbf{m} = l\} = \bigcup_{\pi \in \text{Gr}(l, r)(\mathbf{Q})} \{r\text{-tuples of vectors from } (\pi \cap \mathbf{Z}^r)^d \text{ with rank } l\}.$$

In fact, the sets whose union we are taking constitute an $\text{SL}(d, \mathbf{Z})$ -invariant partition. Thus we need to parametrize linearly independent vectors in $(\pi \cap \mathbf{Z}^r)^d$ for each π . Let $B: \mathbf{R}^l \rightarrow \mathbf{R}^r$ be a linear map with image π such that $B(\mathbf{Z}^l) = \pi \cap \mathbf{Z}^r$. There can be many of these; any one will do. Using the standard basis, $B = (b_i^j)$ with $1 \leq i \leq l$ and $1 \leq j \leq r$, and we obtain $\mathbf{m}^j = \sum_i b_i^j \mathbf{n}^i$, where $\mathbf{n}^i \in \mathbf{Z}^d$ form a linearly independent set. Thus the integral becomes

$$\frac{1}{\mu(G/\Gamma)} \int_{G/\Gamma} \sum_{\pi \in \text{Gr}(l, r)(\mathbf{Q})} \left[\sum_{\substack{\mathbf{n}^1, \dots, \mathbf{n}^l \in \mathbf{Z}^d \\ \text{linearly indep.}}} F(\mathbf{M} \sum b_i^1 \mathbf{n}^i, \dots, \mathbf{M} \sum b_i^r \mathbf{n}^i) \right] d\mu(\mathbf{M}).$$

The quantity in brackets is Γ -invariant, so the sum over π can be interchanged with the integral.

For each π and B we can now apply the Lemma. It gives

$$\sum_{\pi \in \text{Gr}(l, r)(\mathbf{Q})} \int_{(\mathbf{R}^d)^r} F(\sum b_i^1 \mathbf{x}^i, \dots, \sum b_i^r \mathbf{x}^i) d\mathbf{x}^1 \dots d\mathbf{x}^l.$$

Since $B(\mathbf{Z}^l) = \pi \cap \mathbf{Z}^r$, the Jacobian of B is the covolume of π . The statement of the Theorem follows after a linear change of variables. □

Proposition 7. *For any Ξ and τ we have*

$$\mathbf{P}_{1, \sigma, \Xi, \tau / \Xi}^{(d)} \implies \text{Pois } \sigma$$

as $d \rightarrow \infty$.

Proof. From (2), all we need to show is that moments of $f_{\tau, \sigma}$ are Poissonian for large d . First consider that case when $\xi \notin \mathbf{Q}$. Without loss of generality we set $\tau = 0$. Then we need to find

$$\lim_{d \rightarrow \infty} \int_{\mathbf{M} \in G/\Gamma} \int_{\mathbf{v}} \left[\sum_{\mathbf{m} \in \mathbf{Z}^d \setminus \{0\}} (\chi_1 \dots \chi_l \cdot \chi_\sigma)(\mathbf{M}\mathbf{m} + \mathbf{v}) \right]^k d\mu(\mathbf{M}) d\mathbf{v} \quad (7)$$

for $k = 0, 1, \dots$. Taking the integral over \mathbf{v} inside the sum, we clear the way for Theorem 5 applied to

$$F(\mathbf{x}_1, \dots, \mathbf{x}_d) = G_1(\mathbf{x}_1) \dots G_1(\mathbf{x}_{d-1}) G_\sigma(\mathbf{x}_d)$$

where

$$G_t(\mathbf{z}) = \int_{y=0}^1 \chi_t(z_1 + y) \dots \chi_t(z_k + y) dy.$$

For any plane π' as in the Theorem, we have that

$$\int_{\pi'} F(x) dx / (\text{covol } \pi)^d = \left(\frac{\int_{\pi} G_1(\mathbf{x}_1) d\mathbf{x}_1}{\text{covol } \pi} \right)^d \cdot \frac{\int G_{\sigma}(\mathbf{x}_d) d\mathbf{x}_d}{\int G_1(\mathbf{x}_d) d\mathbf{x}_d}. \quad (8)$$

It is elementary to see that the quantity raised to the power d is at most one: the numerator is the volume of π “lying” inside the “crystal” shape, and $\text{covol } \pi_{\mathbf{Z}}$ is the volume of a fundamental domain. To wit, consider first the case when $G_t(\mathbf{z})$ is replaced by the indicator of $[0, 1]^k$. Since vertices of any fundamental domain for $\pi_{\mathbf{Z}}$ have integer coordinates, it can completely cover the part of the plane inside the cube. Furthermore, the quantity in parentheses can equal one only when $\pi \cap [0, 1]^k$ constitutes a fundamental domain for $\pi_{\mathbf{Z}}$. This means that there exists a \mathbf{Z} -basis $\{e_i\}_1^k$ for $\pi_{\mathbf{Z}}$ such that

- $e_i \in \{0, 1\}^k$, $1 \leq i \leq k$;
- $e_i + e_{i'} \in \{0, 1\}^k$, $1 \leq i, i' \leq k$.

Hence two distinct e_i , $e_{i'}$ cannot take on the value 1 in the same coordinate. The same argument extends to other cubes of the form $[-s, 1-s]^k$ for $s \in [0, 1]$ and so, too, for the original $G_t(\mathbf{z})$ as it is an average over cubes of this kind.

The above argument shows that the limit as $d \rightarrow \infty$ exists for each moment and that rate of convergence is exponential. To understand this limit, we focus on the terms with $\int_{\pi} G(x_1) dx_1 / \text{covol } \pi = 1$. Since in (7) we omit the terms in which any of $m^l = 0$, the only terms that survive after taking the limit are the ones with π generated by e_i for which $\sum_{i=1}^k e_i = (1, \dots, 1)$ (no zero coordinates). For planes π of fixed dimension l the number of possibilities is the number of partitions of a set of k elements into l non-empty subsets, which is exactly $S(k, l)$. Finally observing that the last factor in (8) is $\sigma^{\dim \pi} = \sigma^l$, we find that the k -th moment tends to

$$\sum_{l=1}^k S(k, l) \sigma^l,$$

which is the corresponding moment of the Poisson distribution with parameter σ .

In the case when $\xi = p/q \in \mathbf{Q}$ we can modify the above proof. The integral over v becomes a finite sum, and we let $G_t(z) = \frac{1}{q} \sum_{r=0}^{q-1} \chi_t(z_1 + r/q) \dots \chi_t(z_d + r/q)$; the variable τ appears in an equation similar to (8) and doesn't enter the definition of $G_t(z)$. The statements from the continuous version are true for this function as well (since it is also an average over cubes), and the proof is complete. □

Generalizing this proposition we can obtain the statement of Theorem 2.

Proof of Theorem 2. What we need to show is that

$$\int_{\mathbf{M}} \int_{V \in \Xi^d} \sum_{\substack{\mathbf{m}^{1,1}, \dots, \mathbf{m}^{k^1,1} \\ \vdots \\ \mathbf{m}^{1,n}, \dots, \mathbf{m}^{k^n,n} \\ \in \mathbf{Z}^d \setminus \{0\}}} \prod_{j=1}^n \prod_{j'=1}^{k^j} \chi_1 \cdots \chi_1 \cdot \chi_{(\tau^j, \sigma^j + \tau^j)}(\mathbf{M} \mathbf{m}^{j',j} + \mathbf{v}^j) d\mu(\mathbf{M}) dV$$

has a limit as $d \rightarrow \infty$ for every choice of k^1, \dots, k^j . If $\mathbf{Y}_{n,\sigma,\Xi,\tau/\Xi}^{(d)} = (Y^1, \dots, Y^n)$ is distributed according to $\mathbf{P}_{n,\sigma,\Xi,\tau/\Xi}^{(d)}$, then this expression is nothing more than the moment of order (k^1, \dots, k^j) .

Now we make a simplifying observation: we can assume that $k^j = 1$ for all j without loss of generality since taking all possible n and Ξ and computing $\mathbf{E} \prod Y^j$ produces all the moments $\mathbf{E} \prod (Y^j)^{k^j}$. That is, duplicating the random variable Y^j k^j times allows us to assume that k^j is 1. So we need to analyze

$$\int_{\mathbf{M}} \int_{V \in \Xi^d} \sum_{\mathbf{m}^j \in \mathbf{Z}^d \setminus \{0\}} \prod_{j=1}^n \chi_1 \cdots \chi_1 \cdot \chi_{(\tau^j, \sigma^j + \tau^j)}(\mathbf{M} \mathbf{m}^j + \mathbf{v}^j) d\mu(\mathbf{M}) dV,$$

which by Theorem 5 is

$$\sum'_{\pi \in \text{Gr}(r,l)(\mathbf{Q})} \int_{V \in \Xi^d} \int_{\pi'} \frac{dx}{(\text{covol } \pi_{\mathbf{Z}})^d} \prod_{j=1}^n \chi_1 \cdots \chi_{(\tau^j, \tau^j + \sigma^j)}(x^j + \mathbf{v}^j) dV. \quad (9)$$

Since \mathbf{m}^j are non-zero, we exclude the “coordinate planes” as in Remark 3; this is denoted by the prime in the formula above.

We need to account for planes $\pi \in \text{Gr}(r,l)(\mathbf{Q})$ that will contribute in the limit $d \rightarrow \infty$. By the argument from the previous proposition we have that

$$\int_{\pi'} \prod_{j=1}^n \chi_1 \cdots \chi_{(\tau^j, \tau^j + \sigma^j)}(x^j + \mathbf{v}^j) dx \leq (\text{covol } \pi_{\mathbf{Z}})^d. \quad (10)$$

Since Ξ^d is normalized to have measure 1, it suffices to study the integrand for fixed $V \in \Xi^d$. If we can find V and π for which strict inequality is true in (10), then by continuity we have strict inequality for the integral over $V \in \Xi^d$ and thus conclude that π doesn't contribute in the limit. We will do this for $V = 0$ first. A plane that will contribute in the limit $d \rightarrow \infty$ must satisfy the property that $\pi \cap [0, 1]^r$ is a fundamental domain for $\pi \cap \mathbf{Z}^r$ as in the previous proposition. For each of these planes we can try to find another V that gives strict inequality in (10). If $V \in \pi$, we are translating the cube along the plane and thus getting the same cross-sectional area. So suppose $V \in \Xi \setminus \pi$; this corresponds to cutting the cube with a plane parallel to π . It is easy to see that for such planes the section will always have smaller area than the one for $V \in \pi$. Thus it must be the case that $\Xi \subset \pi$.

To summarize, a plane π contributes to the limit only if $\pi \cap [0, 1]^r$ is a fundamental domain for $\pi \cap \mathbf{Z}^r$ and $\Xi \subset \pi$. This means that $\mathbf{P}_{n,\sigma,\Xi,\tau/\Xi}^{(d)}$ has a limit as $d \rightarrow \infty$ because all moments

exist. If we write $\xi = (\xi^1, \dots, \xi^1, \xi^2, \dots, \xi^2, \dots, \xi^{n'}, \dots, \xi^{n'})$ reordering as necessary, then π must be a product of admissible planes for (ξ^1, \dots, ξ^1) , (ξ^2, \dots, ξ^2) , \dots , $(\xi^{n'}, \dots, \xi^{n'})$. Hence the moment will split as the product of moments over distinct ξ^j . Using this observation and the previous proposition we see that in the case of distinct ξ^j in Theorem 2 the limiting distribution is the product of independent Poisson distributions. If $\xi^j = \xi^{j'}$ but $(\tau^j, \tau^j + \sigma^j) \cap (\tau^{j'}, \tau^{j'} + \sigma^{j'}) = \emptyset$, then such ξ^j and $\xi^{j'}$ behave as if they were unequal since the factor

$$\prod \chi_{(\tau^j, \tau^j + \sigma^j)}(x^j + \mathbf{v}^j)$$

from (10) vanishes. It is evident that if $(\tau^j, \tau^j + \sigma^j) \cap (\tau^{j'}, \tau^{j'} + \sigma^{j'}) \neq \emptyset$ for some j, j' , then the limiting distribution cannot be a product of independent distributions. This concludes the proof of Theorem 2. \square

References

- [1] Jens Marklof. The n -point correlations between values of a linear form. *Ergodic Theory and Dynamical Systems*, 20(4):1127–1172, 2000. With an appendix by Zeév Rudnick.
- [2] Jens Marklof. Distribution modulo one and Ratner’s theorem. In *Equidistribution in number theory, an introduction*, volume 237 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 217–244. Springer, Dordrecht, 2007.
- [3] Calvin C Moore. Ergodicity of flows on homogeneous spaces. *American Journal of Mathematics*, 88:154–178, 1966.
- [4] Marina Ratner. On Raghunathan’s measure conjecture. *Annals of Mathematics. Second Series*, 134(3):545–607, 1991.
- [5] C. A Rogers. The moments of the number of points of a lattice in a bounded set. *Philosophical Transactions of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 248:225–251, 1955.
- [6] Nimish A Shah. Limit distributions of expanding translates of certain orbits on homogeneous spaces. *Indian Academy of Sciences. Proceedings. Mathematical Sciences*, 106(2):105–125, 1996.